IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-9, Sept. 2018 E-ISSN: 2347-2693

Designing a Graph Anonymization Framework for Secure Packet Transmission in the IP Over Ethernet LAN

R. Ganguli[1*], S. Roy[2]

[1]Dept. of Computer Science, The Bhawanipur Education Society College, Calcutta University, Kolkata, India
[2]Dept. of Computer Science, The Heritage College, Calcutta University, Kolkata, India

Available online at: www.ijcseonline.org

Accepted: 19/Sept/2018, Published: 30/Sept/2018

Abstract— Ethernet Local Area Network (LAN) has become almost omnipresent today. Ethernet has scored over its competitor LAN technologies because of its simplicity, ruggedness and low cost as well as, the advantages it enjoyed owing to its early entry into the LAN market. Despite many research work and development efforts in the area of data communication's security, the importance of internal local area network LAN security is still underestimated. Ethernet LAN has always been known to be an insecure technology. There is no way in the Ethernet LAN protocol to check the authentication or message integrity. Intruder can hack the IP packet header and create packets with false IP addresses or MAC addresses. In this paper we discuss about the design of a secured packet transport scheme in the IP over Ethernet LAN (IPEL) based on graph anonymization approach. To ensure secure communication, the identity of IP packet source and destination needs to be hidden from intruder. In our proposed scheme each host in the IPEL encrypts the header part of IP packet using graph anonymization algorithm and forwards to the next node. On receiving the packet host performs decryption using graph de-anonymization algorithm to read the packet header. This approach of securing IP packet header offers good scalability, possibly low cost and also promises a reasonable level of security in the IPEL.

Keywords—Encryption, Decryption, IPEL, IPSec, SSL, Graph Anonymization

I. INTRODUCTION

Early Internet was developed at the United States Defense Advanced Research Projects Agency (DARPA or ARPA)[1]. This worked fine at first when there were only a few networks on the Internet and security was not a major concerned at that time. At the early stage of the Internet, it was used primarily for researchers and other networking professionals [2][4]. However, with the expansion of the Internet, when it was used for commercial purpose, maintaining the security of the network becomes a major concern for the network administrator [1][2]. Protection of data during its entire transit path from the source host to the destination host gains importance due to the exponential growth of the complex network structure including all kinds of hardware and software technologies involved in the data communication. Network Security is of utmost importance in the Ethernet LANs. Confidentiality, integrity and authentication are key issues in data protection [1][3]. Confidentiality or privacy is the key term used to prevent the disclosure of information to unauthorized individuals or systems. The system attempts to enforce confidentiality by encrypting the message. Confidentiality is necessary for maintaining the privacy of the system. Integrity means that the data cannot be altered and must arrive at the receiver as same as they were sent. Integrity is violated when a message is modified in its transit path.

Although there exists various security threats which include Denial of Service (DoS) attack (Flooding the network to prevent legitimate network traffic), Eavesdropping (Act of intercepting communications between two terminal hosts), Phishing (Gaining sensitive information like login password and other confidential information like credit card and debit card information) or IP Spoofing [3], which is our major concern in this paper. IP spoofing is an attack where an intruder gains unauthorized access to the header part of IP packet. Using IP Spoofing an attacker impersonates as another host by manipulating IP packets and it seems to the receiver as if the message has come from a trusted host. In order to impersonate as a trusty one, the intruder first learns the IP address of a trusted system, and then replaces the packet header with a spoofed source IP address. Due to the promiscuous mode of operation in Ethernet LANs [4][5], any intruder lying in the transit path between the source host and the destination host can sniff a packet, can read the contents of the data packet (loss of privacy), can change the content of the packet (loss of integrity) or can pretend as another host (loss of authentication) called IP spoofing[6][7]. In section I, we discussed the present security scenario of an IP packet in

Ethernet LAN. Section II summarizes the related work explored in this area of securing IP packet transmission in Ethernet LAN. The concept of graph ammonization and its application are discussed in Section III. Section IV mainly focuses on the basic design approach and working methodology of graph anonymization scheme in message hiding. The paper concludes in section V mainly pointing out the future scope of the work.

II. RELATED WORK

A number of methods have been deployed over the past few years to address the need for security. Most of these are focused at the higher layers of the TCP/IP stack in order to compensate for IP's lack of security [8][9]. For example, certain applications like World Wide Web access or File Transfer Protocol (FTP) uses Secure Sockets Layer (SSL) protocol. Simply SSL is all about maintaining secure connections in a web. Netscape Communications Corp established the secure socket layer (SSL) which was introduced into a new layer between the application layer and the transport layer to enhance security[9][10]. Main functionality in this layer is to compress and encrypt data. In addition, it also automatically detects whether the data has been tampered in transit. SSL is basically used in web browsers, but it can also be used in other applications as well in conjunction with HTTP called HTTPs which is an application layer protocol.

Since IP datagram packets are usually routed between two hosts over the public Internet, any information in the packet can be hacked by an intruder. With the growing use of the Internet for critical applications, security enhancements are needed also at the network layer.

The IETF has established a set of protocols that provide a secure Internet connection, known as IPSec (IP Security)[10][11]. IPSec offers authentication and privacy services at the network layer. IPSec provides a general framework for each pair of communicating hosts to select encryption[10] algorithm and key length. IPSec is a set of services and protocols that provide a complete security solution for an IP network. Since IPSec works at the IP layer, it can provide these protections for any higher-layer TCP/IP application or protocol without the need for additional security methods, which is a major strength. IPSec does not rely on a single algorithm, rather depends on multiple algorithms so that failure of a single algorithm does not have any effect on the overall security. IPSec requires each receiver to gather all details about a security scheme into an abstraction known as a Security Association (SA) which involves security identifiers. IPSec can operate in two main modes, one is Transport Mode and the other is Tunnel Mode[10]. In the transport mode, a header which includes the SA identifier, sequence number and other security information is attached after the IP header. In the tunnel mode, IP packet along with its header part is encapsulated to form a new IP packet with a new IP header. The major advantage of IPSec is in its transparency and independency of upper layer protocols. Since IPSec operates at Layer 3, it is indifferent as to whether application traffic is being transported using TCP or UDP protocols. In fact, IPSec, with conjunctions to other cryptographic network protocols such as Transport Layer Security (TLS) which works as the transport layer, Secure Shell and HTTPS which work at the application layer is used to prevent IP Spoofing. Despite its tremendous advantages IPSec also suffers from certain limitations. To run IPSec properly, we need IPSec computability software running on the attached hosts. IPSec has a drawback where any flaw existing at the IP layer in the remote network get passed to the business network through the IPSec tunnel. Moreover, IPSec requires VPN enabled setup at both ends for a secure communication channel [11]. Since IPSec does not have any standards or fixed encryption algorithm, weaker encryption or incorrect deployment makes IPSec much less efficient than SSL. For two hosts to communicate via an IPSec connection, implementation gets complex since both must agree on the same security policy, called a security association, which must be configured in the hosts on both ends of the IPSec connection.

From the above discussion, it is clear that IPSec is one of the preventive measures of network layer protocol used for protecting IP packet header from spoofing. But, in spite of its wide use, it has certain drawbacks like complex technology, Virtual Private Network (VPN) enabled end terminal etc. In our proposed approach we have introduced an alternative graph based encryption called graph anonymization to protect the IP header. The security schemes we have used have low complexity, simple encryption/decryption algorithm and possibly low cost. Degree anonymity is NP-hard on 3-colorable graph [12][13] .Thus using this technique to encrypt IP packet header ensures high level of security from any attack in IPEL.

III. GRAPH ANNONYMIZATION

In this section we formalize our definition of graph anonymity. Graph Anonymization problem states that for given a graph *G*, there exists a *k*-degree anonymous graph that stems from *G* with minimum number of graph-modification operations. This graph modification operation includes edge additions/ edge deletions, node addition/deletion.

A graph is called *k*-degree anonymous if for every node *v*, there exist at least *k*-1 other nodes in the graph with the same degree as *v*.

Definition: Let $G(V,E)$ be a simple graph; V is a set of nodes and E the set of edges in G. We use d(G) to denote the *degree sequence* of G. That is, d(G) is a vector of size $n = |V|$ such that d$G(i)$ is the degree of the i-th node of G.
Without loss of generality, we also assume that entries in d are ordered in decreasing order of the degrees they correspond to, that is,
d(1) >= d(2) >= ,.....,>=d(n).

Definition 1: *A vector of integers* v *is k-anonymous, if every distinct value in* v *appears at least k times.*
For example, vector v = [5,5,3,3,2,2] is 2-anonymous.

Definition 2: *A graph $G(V,E)$ is k-degree anonymous if the degree sequence of G,* d(G)*, is k-anonymous.*
Figure 1 shows two examples of degree-anonymous graphs. In figure 1(a), all three nodes have the same degree and thus the graph is 3-degree anonymous. Similarly, in figure 1(b), the graph is 2-degree anonymous having two nodes with degree 1 and four nodes with degree 2.

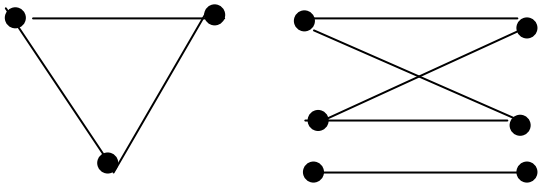

Figure 1: Example of a (a) 3-degree anonymous graph (b) 2-degree anonymous graph

Here, definition for graph anonymity is inspired by the notion of k-anonymity for each IP packet. These IP packets identifiable by its public attributes are required to be hidden in a group of size k. The goal of an anonymization scheme is to prevent such an intruder from uniquely identifying each packet header in the anonymized graph.

More formally we define the (k, ℓ)-anonymity property of a graph where for every vertex in the graph there exist at least k other vertices that share at least ℓ of its neighbors. In order to meet this anonymity requirement one could transform any graph into a complete graph[12] which although preserves privacy but such a graph becomes useless for studying.

With the tremendous growth of large internetworking, flow of IP packets increase through the public network, and subsequently the privacy protection of the IP header part becomes a major demand. In a landmark paper, Liu and Terzi [14] considered the vertex degrees as feature; see Wu et al. [14] for other features considered in the literature. Therein, by the definition of k-anonymity of a graph, different values of k reflect different privacy demands and the natural computational task arises. Given some fixed value k, we need to make few changes to the graph in order to make it k-anonymous. Liu and Terzi [4] proposed a heuristic algorithm where a graph is made k-anonymous by edge addition. In this paper, we consider minimum edge addition on the underlying graph of interactions i.e. IP packets through open unsecured IPEL, while preserving the privacy of the packet headers. More specifically, the private information i.e., the source and destination IP address are to be protected from any intruder lying in the IPEL. Therefore, we design an anonymization framework that tries to hide the identity of packet by creating groups of packets. We call such packets anonymized. Our goal is to anonymize all packets of the graph by introducing minimal changes to the overall graph structure.

IV. BASIC DESIGN AND METHODOLOGY USING GRAPH ANNONYMIZATION

The problem here is concerned with taking a message and dividing it into block of 8 characters and represents each of these blocks as separate graphs. The same anonymization algorithm is applied on each of these graphs. These graphs after anonymization can be sent to the next node i.e. the receiver. On the receiving end, the graphs are de-anonymized one by one in order they are received. The receiver uses the de-anonymization algorithm which is just the reverse of that of the anonymization algorithm. But the message extracted will be meaningful only when all the graphs received are de-anonymized. Then only the original message is retrieved.

The message to be encrypted from the IP packet is the header section only. For simplicity we consider only source and destination IP address which always comes in bytes, i.e. length is multiples of 8. Thus the header length justifies the block size taken in the algorithm.
The length of the original message, i.e. length of the IP packet header in this case is known to all the nodes as they are agreeing on the same protocol. This helps the receiver to apply the de-anonymization algorithm to get back the original message.
Graphs here are represented in the form of adjacency matrix. The anonymization and de-anonymization algorithm are applied on these adjacency matrices and the degree sequence obtained from the matrices.
Graph anonymization is achieved by applying edge addition operation on the graphs. We have taken undirected, unweighted graphs for this problem. The degree sequence are changed into k-degree anonymous degree sequence by adjusting and the corresponding changes in the graph makes the graph k-degree anonymous.
The following steps collectively ensure encrypting of the header part of IP packet to avoid IP Spoofing in the IPEL. It is to be noted that we ensure security in the transit path i.e. IPEL only.

1. The existing TCP/IP Protocol suit is modified and an additional layer is introduced between IP layer and the Transport Layer in the TCP/IP stack.
2. The algorithm will be implicitly present in the TCP/IP protocol stack of each host in the IPEL.
3. As soon as a host sends a packet it first encrypts the IP header using graph anonymization algorithm. Encrypted packet is then sent to the desired destination host in the IPEL.
4. Due to promiscuous mode of Ethernet LAN, each host connected in the IPEL picks up the packet and decrypts the header using the de-anonymization algorithm to see destination IP address. If the destination address does not match it simply ignores the packet. Only the intended receiving host will accept the packet after decrypting. Figure 2 shows the framework of our design scheme.
5. Here, the key used in encryption-decryption is a status array. It is to be noted that the key (status array) will be generated by the host at run time for packet transmission and same key will be used for all hosts in the IPEL in the entire session. A standard Diffie-Hellman key exchange algorithm is used to exchange the key between any two hosts in the IPEL.

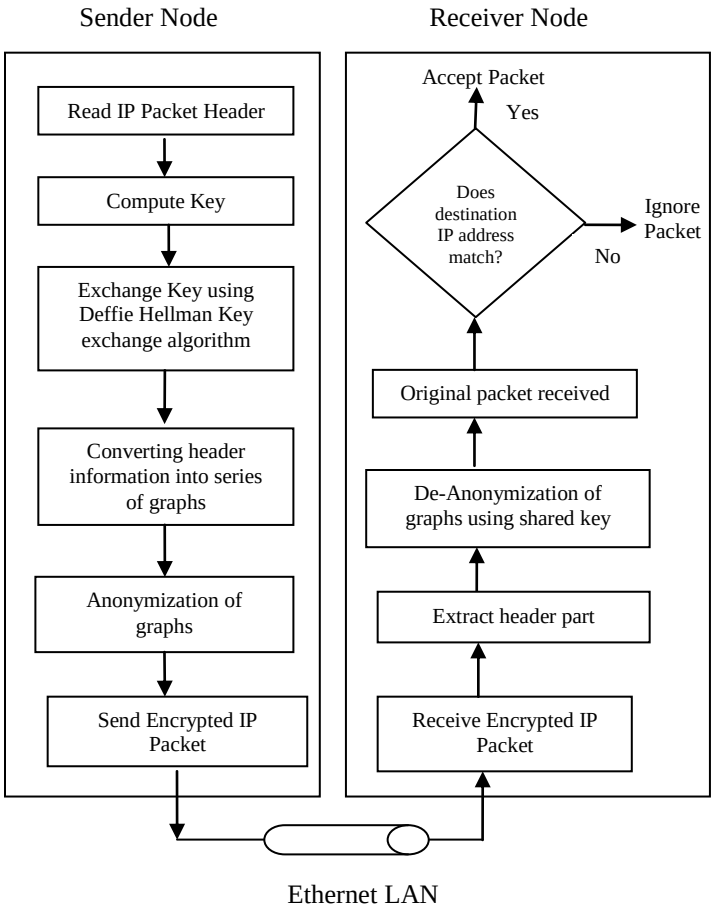

Figure 2: Proposed Anonymized Architecture in the IPEL

V. CONCLUSION AND FUTURE SCOPE

This paper has mainly aimed to prevent incorrect packet forwarding in the IPEL. We have proposed a framework using the concept of graph anonymization. This method is less costly and simpler in design to prevent from IP/MAC spoofing attack. In our scheme, each host, in the IPEL encrypts the IP packet header.. The algorithm has used only edge addition operation on the graphs. Whereas the effect of using vertex addition/deletion has not been studied which leaves ample scope for further exploration. Also the block size into which the message is broken is fixed here. Here, our proposed algorithm always makes the graphs complete while anonymizing. Complete property of a graph mostly confuses adversary as the graph remains no longer useful for analysis. However extending this concept from the Ethernet LAN to internetwork (consisting of subnet and routers) is yet to be studied. Further simulation of this algorithm to achieve optimized results, i.e., adding minimum number of edges in the graph, is the future scope of our work.

REFERENCES

[1] A.S. Tenenbaum,"ComputerNetworks", 4th Ed., Pearson Education Asea, LPE, 2003.
[2] J.F. Kurose ad K.W. Ross, "Computer Networking: A Top-Down Approach Featuring the Internet", 3rd Ed., Pearson Education Asea, LPE,2005.
[3] B. A. Forouzan, "Data Communications ad Networking", 4th Ed., Tata McGraw-Hill, New Delhi, 2004.
[4] Metcalfe, R.M ad Boggs, D.R, "Ethernet: Distributed Packet Switching for Local Computer Networks", Communication of the ACM. Vol 19, pp 395-404, July 1976.
[5] C.E. Spurgeon,, "Ethernet-The Definitive Guide", Orielly/Shroff Publishers & Distributors (India), 2000.
[6] A Leon,-Garcia and I. Widjaja, "Communication Networks", 2nd Ed., Tata McGraw-Hill, New Delhi, 2004.
[7] R. Perlman, "Interconnections: Bridges ad Routers", Addison Wesley,1994.
[8] L. L. Peterson ad B.S. Davie, "Computer Networks: A systems Approach", 3rd Ed. Morgan Kaufman, 2003
[9] D. E. Comer, "Internetworking with TCP/IP Principles, Protocols, and Architecture", 4th Ed., Prentice-Hall, 2003
[10] W. Stallings, "Network Security Essentials: Application ad Stadards ", 4th Ed.,Pearson Eduction Asea,LPE,2013
[11] J. Katz,Y. Lindell, "Introduction to modern cryptography" Chapman & Hall/CRC Press, 2007.
[12] N.Deo, ,"Graph Theory with Applications to Engineering and Computer Science", New Ed, Prentice-Hall, 2003.
[13] C.C. Aggarwal, Y.Li and P.S.Yu, "On the hardness of graph annonymization." In Proceedings of the 11th IEEE International Conference on Data Mining(ICDN'11),pp 1002-10007,IEEE,2011
[14] K. Liu, , and E. Terzi,. "Towards identity anonymization on graphs". In SIGMOD Conference (2008), pp. 93–106.

Authors Profile

Ms. Runa Ganguli completed her Masters of Science in Computer and Information Science from University of Calcutta, India in the year 2014. She did her Bachelors of Science in Computer Science Honours from Asutosh College, University of Calcutta in 2012. She holds third position in the university in her undergraduate level. She is currently pursuing her Master of Technology in Computer Engineering and Applications from A.K. Chowdhury School of IT, University of Calcutta. She is currently working as Assistant Professor in the Department of Computer Science, The Bhawanipur Education Society College, Kolkata, University of Calcutta, India since 2015. Earlier she worked in Asutosh College as Lecturer for 1 year. She has several papers in reputed peer reviewed journals and international conferences. Her main research interest includes Graph Theory & its Applications,Graph Based Database, Network Security and Software Engineering. She has 4 years of teaching experience.

Dr. Siddhartha Roy completed his PhD in Engineering from Jadavpur University, West Bengal, India. He has completed his 1st Master Degree (MCA) from Indian Institute of Engineering Science and Technology, Shibpur, West Bengal, India. He completed his second Master degree MBA with specialization in operation management from IGNOU, India.. He did his graduaton in Mathematics Honours from St Xavier's College,University of Calcutta. He worked in the software industry for one year and after taking professional development training from TAFE, Adelaide, South Australia, came to the academic world. At present he is an Assistant Professor in the Computer Science Department of The Heritage College, Kolkata. He has several publications both in national and international level peer-reviewed journal. He is also an author of text books in undergraduate levels. . His main research work focuses on Routing Algorithms, Network Security, Data Mining, and Bioinformatics based education. He has 17 years of teaching experience and 5 years of Research Experience.

